

Systematic Gene Search in the Incyte LifeSeq Database

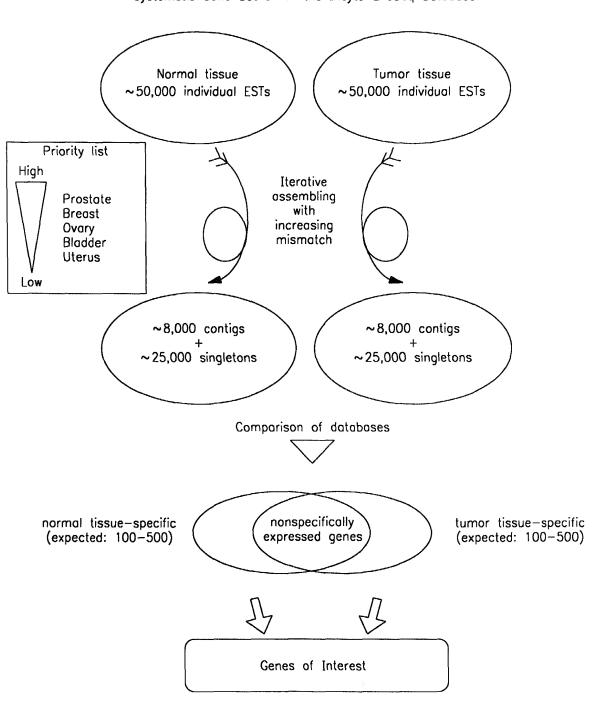
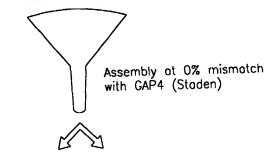


FIG. I

~50,000 ESTs per tissue



Contigs Singletons

Contigs increasing in number and length

Iterative assembly with increasing mismatch (1%, 2%, 4%)

5000-6000 Contigs

~25,000 other singletons



~30,000 consensussequences per tissue

FIG. 2a

8-38 Stor. 87 St

×4., (

Harry Marie Street King of the Control of the Contr

~50,000 ESTs of a tissue (e.g.: uterus tumor) GAP4 Assembly 1st Round: minimum initial match: 20 maximum pads per read: 8 maximum percent mismatch: 0 GAP4-Database 1: unassembled **ESTs** Contigs 1 Singletons 1 GAP4 Assembly 2nd Round: minimum initial match: 20 maximum pads per read: 8 maximum percent mismatch: 1 GAP4-Database 2: unassembled **ESTs** Contigs 2 Singletons 2 GAP4 Assembly 3rd Round: minimum initial match: 20 maximum pads per read: 8 maximum percent mismatch: 2 GAP4-Database 3: unassembled **ESTs** Contigs 3 Singletons 3

FIG. 2b-I

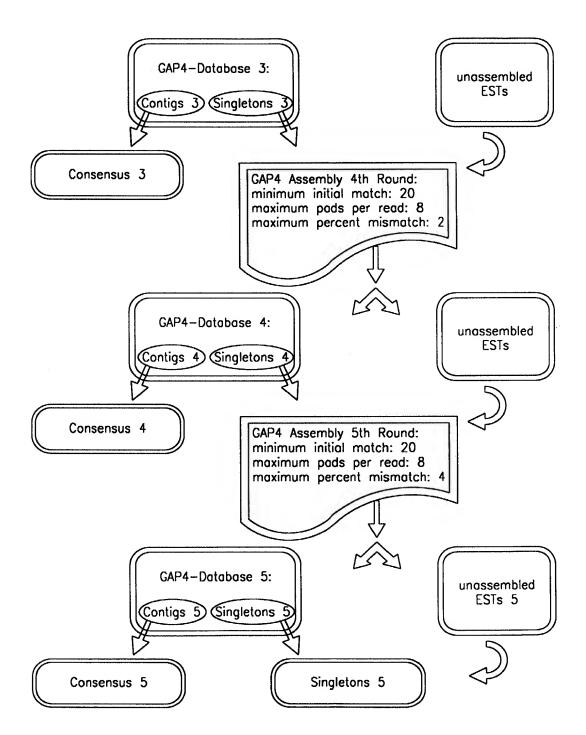


FIG. 2b-2

A control of the cont

Şadı

FIG. 2b-3

West of the first feet of the same feet

The state of the s

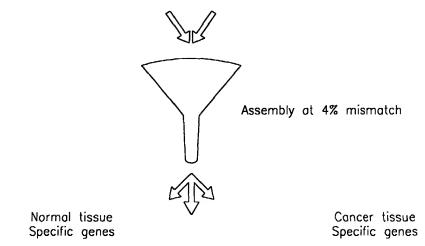
FIG. 2b-4

DRAFTSMAN

In silico subtraction of gene expression in various tissues

~30,000 consensus sequences normal tissue

 $\sim 30,000$ consensus sequences cancer tissue



Genes expressed in both tissues

FIG. 3

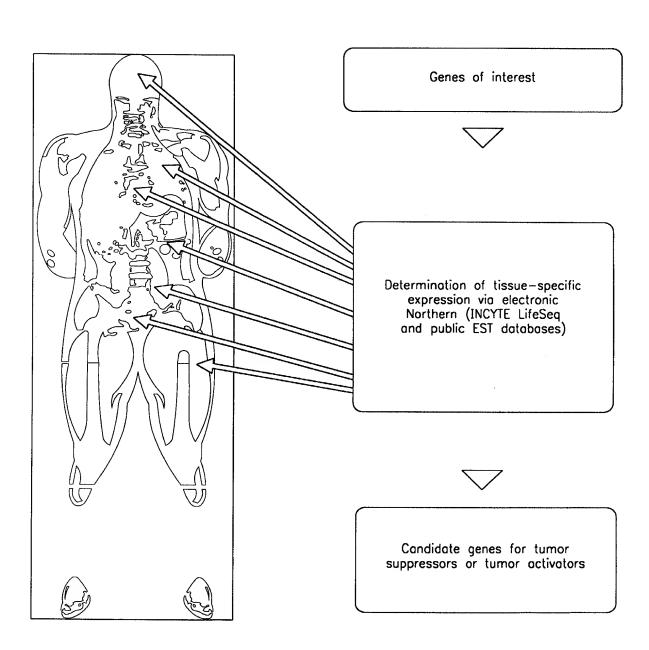


FIG. 4a

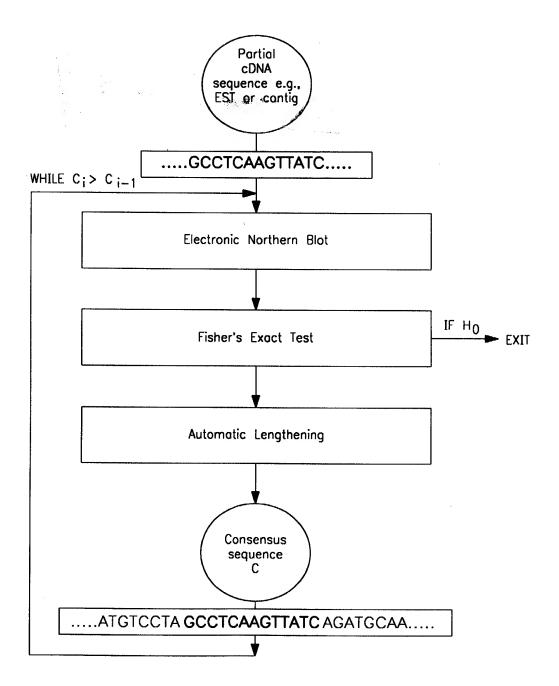


FIG. 4b

Isolation of genomic BAC and PAC clones

Chromosomal clone localization via FISH

Hybridization signal



Sequencing of clones that are located in regions that have chromosomal deletions in prostate and breast cancer leads to identification of candidate genes

Exon Intron

Confirmation of candidate genes by screening of mutations and/or deletions in cancer tissues

FIG. 5